UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 09-04
Statistics and Econometrics Series 03
January 2009

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

WAVELET-BASED DETECTION OF OUTLIERS IN VOLATILITY MODELS

Aurea Grané(1) and Helena Veiga(2)

Abstract

Outliers in financial data can lead to model parameter estimation biases, invalid inferences and poor volatility forecasts. Therefore, their detection and correction should be taken seriously when modeling financial data. This paper focuses on these issues and proposes a general detection and correction method based on wavelets that can be applied to a large class of volatility models. The effectiveness of our proposal is tested by an intensive Monte Carlo study for six well known volatility models and compared to alternative proposals in the literature, before applying it to three daily stock market indexes. The Monte Carlo experiments show that our method is both very effective in detecting isolated outliers and outlier patches and much more reliable than other wavelet-based procedures since it detects a significant smaller number of false outliers.

Keywords: Outliers, Outlier Patches, Volatility Models, Wavelets.

JEL classification: C22, C5.

(1) A. Grané, Statistics Department, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), Spain, e-mail: aurea.grane@uc3m.es . Corresponding author.
(2) H Veiga, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), Spain; ISCTE Business School, Avenida das Forças Armadas, 1600-083 Lisboa, Portugal, e-mail: mhveiga@est-econ.uc3m.es
Work partially supported by Spanish grants SEJ2006-03919, SEJ2007-64500 and MTM2006-09920 (Ministry of Education and Science- FEDER).

Wavelet-based detection of outliers in volatility models

Aurea Grané(1) Helena Veiga(2)

(1),(2) Statistics Department. Universidad Carlos III de Madrid.
(2) Finance Research Center. ISCTE Business School.

Abstract

Outliers in financial data can lead to model parameter estimation biases, invalid inferences and poor volatility forecasts. Therefore, their detection and correction should be taken seriously when modeling financial data. This paper focuses on these issues and proposes a general detection and correction method based on wavelets that can be applied to a large class of volatility models. The effectiveness of our proposal is tested by an intensive Monte Carlo study for six well known volatility models and compared to alternative proposals in the literature, before applying it to three daily stock market indexes. The Monte Carlo experiments show that our method is both very effective in detecting isolated outliers and outlier patches and much more reliable than other wavelet-based procedures since it detects a significant smaller number of false outliers.

Keywords: Outliers, Outlier Patches, Volatility Models, Wavelets.

JEL classification: C22, C5

1 Introduction

Financial time series typically exhibit excess of kurtosis and volatility clustering, which consists of periods of high (low) volatility followed by periods of high (low) volatility. Several models had been proposed in the literature with the aim to capture these features. The ARCH model by Engle (1982) and the GARCH model by Bollerslev (1986) became benchmarks models in finance, specially due to their easy applicability and flexibility in allowing for simple extensions that better fit the empirical facts of financial data. Indeed, since the estimated residuals computed from the GARCH model often have excess of kurtosis, Bollerslev (1987) introduced

(1) She acknowledges financial support from the Spanish Ministry of Education and Science and FEDER, research projects MTM2006-09920 and SEJ2007-64500. Corresponding author.
(2) She acknowledges financial support from the Spanish Ministry of Education and Science, research project SEJ2006-03919.
Authors' address: (1),(2) Statistics Department, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe, Spain. (2) ISCTE Business School, Avenida das Forças Armadas, 1600-083 Lisboa, Portugal.
E–mail: A. Grané, `aurea.grane@uc3m.es`; H. Veiga, `mhveiga@est-econ.uc3m.es`.
Date: January 26, 2009.

a t-distributed GARCH model by allowing the error term to follow a Student's t distribution. This slight modification allows the model to reach levels of kurtosis more comparable to the ones observed in the data. However, it can be observed that the estimated standardized residuals from this extension still register excess of kurtosis (see Baillie and Bollerslev, 1989; Teräsvirta, 1996). One possible reason for this to occur is that some observations on returns, which are called additive outliers (AO), are not fitted by a gaussian GARCH model and even by a t-distributed GARCH model. The additive outliers can be level outliers (ALO) in the sense that they have effects on the level of the series but not on the evolution of the underlying volatility or volatility outliers (AVO) (see Hotta and Tsay, 1998; Sakata and White, 1998). This last type of additive outliers also affects the conditional variance. Neglecting the existence of these outliers leads to biased parameter estimates (see for example Fox, 1972; Van Dijk et al., 1999), undesirable effects on the tests of conditional homoskedasticity (see Carnero et al., 2007) and to biased out-of-sample forecasts (see for instance Ledolter, 1989; Chen and Miu, 1993a; Franses and Ghijsels, 1999).
This paper focuses mainly on additive (level and volatility) outliers. The effects of innovative outliers on the dynamic properties of the series are less important because they are propagated by the same dynamics like in the rest of the series (see for example Peña, 2001). Our approach is inspired by Bilen and Huzurbazar (2002) who proposed an outlier detection method based on wavelets. Wavelets are a family of basis functions that allows to express and approximate other functions. In fact, wavelet coefficients are able to detect changes in variance, level changes and discontinuities in functions. Hence, they are perfectly suitable for outlier detection. Our method departs from theirs in the way the threshold limits are obtained. They used the proposals suggested by Donoho and Johnstone (1994) and Wang (1995) which rely on the assumption that the data is gaussian. Moreover, since their threshold limits are quite conservative their procedure leads to an extremely high average of false detections. On the contrary, our method for computing the threshold limits is based on the distribution of the maximum of the detailed coefficients (in absolute value) obtained by Monte Carlo. In fact, the threshold is taken to be the 95%-percentile of the distribution of this maximum. In this way, our procedure can be applied to the estimated standardized residuals of different volatility models with errors following any known distribution.
The proposal deals with the estimated model residuals because we are interested in detecting if an "abnormal" observation is an outlier for a particular volatility model. The contributions of this paper are several. First, we propose a method for outlier detection and correction that can be applied to the residuals of different volatility models whose errors may follow any known distribution. Doornik and Ooms (2005) also proposed a procedure to detect additive outliers but it is restricted to GARCH models whose errors follow a normal or a Student's t distribution. Their method is inspired on the work of Chen and Miu (1993b), and therefore, susceptible to the same criticisms. In fact, in order to detect several outliers they apply a recursive procedure. They start by detecting the largest outlier, then they adjust for it and proceed by detecting the second largest outlier and adjust again for it, and so on. In this process, they have to estimate a GARCH model several times and the estimates of the parameters can be affected by the presence of a remaining outlier. On the contrary, our method does not need subsequential estimations of the models and

therefore, it is not susceptible to this criticism. Second, it is well suited for one outlier or multiple outlier detections. Third, it is the one, as far as we know, that detects patches of outliers in different volatility models. Fourth, our detection procedure can be extended to innovative outliers and finally, the method is easy and quick to apply which converts it in an attractive tool to be used by academic communities and/or by practitioners.

Our method is applied to several volatility models, such as: the GARCH, the GJR-GARCH (see Glosten et al., 1993) and the autoregressive stochastic volatility model (ARSV) by Taylor (1986) with errors following a gaussian or a Student's t distribution. The GJR-GARCH is an extension of the GARCH model that allows for an asymmetric response of volatility to changes of market prices. The Monte Carlo results show that our proposal is not only as good as that of Bilen and Huzurbazar (2002) in detecting outliers, whenever both methods can be applied, but also it is much reliable since it detects a significant smaller number of false outliers. This is so, that we may be sure that when it does detect an "abnormal" change in the series analyzed, it is an outlier.

The organization of this paper is as follows. In Section 2 we review the two types of additive outliers introduced by Hotta and Tsay (1998). In Section 3 we make a brief introduction to wavelets and present the algorithm for outlier detection. We study the effectiveness of the proposal through an intensive simulation study and compare it to the method proposed by Bilen and Huzurbazar (2002). In Section 4 we apply it to three daily stock market indexes and then proceed to the correction of the outliers detected. Finally, we conclude in Section 5.

2 Additive outliers in volatility models

The GARH(p, q) models proposed by Bollerslev (1986) and (1987) are given by:

$$y_t = \mu + \varepsilon_t = \mu + \sigma_t \, \epsilon_t,$$

where μ is the conditional mean, ε_t is the prediction error, σ_t^2 is the variance of y_t given information at time $t-1$, $\sigma_t > 0$, $\epsilon_t \sim NID(0,1)$ or follows a Student's t distribution and

$$\sigma_t^2 = \omega + \theta(L)\, \varepsilon_t^2,$$

where

$$\theta(L) = 1 - \frac{\alpha^*(L)}{\beta(L)},$$

with $\omega > 0$, $\alpha^*(L) = 1 - \sum_{i=1}^{q} \alpha_i^* \, L^i$ and $\beta(L) = 1 - \sum_{i=1}^{p} \beta_i \, L^i$, $\beta_i \geq 0$, $\alpha_i^* \geq 0$ and $\sum_{i=1}^{q} \alpha_i^* < 1$ in order to enforce a positive conditional variance and stationarity, respectively.

The conditional variance equation of a GARCH(1,1) given the previous restrictions, can be written as

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

where $\alpha_0 = \omega(1-\beta_1) > 0$, $\alpha_1 = \alpha_1^* - \beta_1 \geq 0$ and $\alpha_1 + \beta_1 < 1$.

The GJR(1,1) model differs from the GARCH(1,1) since it introduces the possibility that positive and negative shocks might affect differently the conditional variance σ_t^2. In fact, the conditional variance equation is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1\, \varepsilon_{t-1}^2\, I_{\{\varepsilon_{t-1}<0\}}(t-1) + \beta_1 \sigma_{t-1}^2,$$

where $I_{\{\varepsilon_t<0\}}(t) = 1$ if $\varepsilon_t < 0$ and 0 otherwise. Moreover, $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and $\gamma_1 \geq 0$ to guarantee a positive conditional variance and $\alpha_1 + \beta_1 + \gamma_1/2 < 1$ to enforce stationary (see Duan et al., 2006). Once more, $\epsilon_t \sim N(0,1)$ or follows a Student's t distribution.

In the context of stochastic volatility, a natural competitor to the GARCH and GJR models is the autoregressive stochastic volatility model (denoted as ARSV(1)) by Taylor (1986). The ARSV model is given by the following expressions:

$$\begin{aligned} y_t &= \mu + \sigma\, \epsilon_t \exp\left(\frac{h_t}{2}\right), & (1) \\ (1-\phi L)\, h_t &= \eta_t. & (2) \end{aligned}$$

In equation (1), μ is the mean of y_t, σ denotes a scale parameter, $\sigma_t = \exp(h_t/2)$ is the volatility of y_t (the return at time t), $\epsilon_t \sim NID(0,1)$ or follows a Student's t distribution and in equation (2), ϕ is the autoregressive parameter, h is an unobserved latent variable that is stationary for $|\phi| < 1$ and $\eta_t \sim NID(0, \sigma_\eta^2)$.

2.1 Additive level outliers (ALO)

The conditional mean equations of the GARCH(1,1) and the GJR(1,1) models with an additive level outlier are defined as:

$$y_t = \mu + \omega_{AO}\, I_T(t) + \varepsilon_t = \mu + \omega_{AO}\, I_T(t) + \sigma_t\, \epsilon_t,$$

where ω_{AO} represents the magnitude of the additive level outlier and $I_T(t) = 1$ for $t \in T$ and 0 otherwise, representing the presence of the outlier at a set of times T. The equations of the conditional variances for the two models remain the same since this type of outlier only affects the level of the series.

In the context of stochastic volatility, the additive level outlier is defined as:

$$\begin{aligned} y_t &= \mu + \omega_{AO}\, I_T(t) + \sigma\, \epsilon_t \exp\left(\frac{h_t}{2}\right), \\ (1-\phi L)\, h_t &= \eta_t, \end{aligned}$$

where ω_{AO} and $I_T(t)$ are defined as before. Examples of additive level outliers may be an institutional change or a market correction that does not affect volatility, among others.

2.2 Additive volatility outliers (AVO)

For this class of outliers, we focus our attention on GARCH-type models, such as the GARCH(1,1) and the GJR(1,1) analyzed before. In this context, an additive volatility outlier for the GARCH(1,1) model is defined as:

$$\begin{aligned} y_t = \mu + \varepsilon_t^* &= \mu + \sigma_t^* \, \epsilon_t, \\ \varepsilon_t^* &= \omega_{AO} \, I_T(t) + \varepsilon_t, \\ \sigma_t^{*2} &= \alpha_0 + \alpha_1 \, \varepsilon_{t-1}^{*2} + \beta_1 \, \sigma_{t-1}^{*2}, \end{aligned} \tag{3}$$

and for the GJR(1,1) model:

$$\begin{aligned} y_t = \mu + \varepsilon_t^* &= \mu + \sigma_t^* \, \epsilon_t, \\ \varepsilon_t^* &= \omega_{AO} \, I_T(t) + \varepsilon_t, \\ \sigma_t^{*2} &= \alpha_0 + \alpha_1 \, \varepsilon_{t-1}^{*2} + \gamma_1 \, \varepsilon_{t-1}^{*2} \, I_{\{\varepsilon_{t-1}<0\}}(t-1) + \beta_1 \, \sigma_{t-1}^{*2}, \end{aligned} \tag{4}$$

where ω_{AO} represents the magnitude of the additive level outlier, $I_T(t) = 1$ for $t \in T$ and 0 otherwise, representing the presence of the outlier at a set of times T as in subsection 2.1 and $I_{\{\varepsilon_t<0\}}(t) = 1$ if $\varepsilon_t < 0$ and 0 otherwise. Note that in both cases, ϵ_t can follow a standard normal distribution or a Student's t distribution.

Finally, we can express σ_t^{*2} in terms of the dynamic effect of the outlier by replacing ε_t^* into equations (3) and (4), respectively. Then for the GARCH(1,1) we have that

$$\sigma_t^{*2} = \alpha_0 + \alpha_1 \, \varepsilon_{t-1}^2 + \beta_1 \, \sigma_{t-1}^{*2} + \alpha_1 \left(2 \, \omega_{AO} \, \varepsilon_{t-1} + \omega_{AO}^2\right) I_T(t-1), \tag{5}$$

and for the GJR(1,1):

$$\begin{aligned} \sigma_t^{*2} = \; & \alpha_0 + \left(\alpha_1 + \gamma_1 \, I_{\{\varepsilon_{t-1}<0\}}(t-1)\right) \varepsilon_{t-1}^2 \\ + \; & \left(\alpha_1 + \gamma_1 \, I_{\{\varepsilon_{t-1}<0\}}(t-1)\right) \left(2 \, \omega_{AO} \, \varepsilon_{t-1} + \omega_{AO}^2\right) I_T(t-1) + \beta_1 \sigma_{t-1}^{*2}. \end{aligned} \tag{6}$$

Therefore, the effect of the AVO outlier affects not only the volatility but also the series level. Its effect in the original series is similar to a patch of ALO outliers with decreasing magnitudes when $\beta_1 < 1$.

Figure 1 shows two daily series with sample size $n = 1000$ generated by a GARCH model with parameters $\{\alpha_0 = 0.0126, \alpha_1 = 0.0757, \beta = 0.9122\}$ in which we forced AVO outliers of size $\omega_{AO} = 15\sigma_y$ at position 213, relatively to the sample size, and $\omega_{AO} = 25\sigma_y$ at position 500, relatively to the sample size, where σ_y is the standard deviation of the series y_t.[1]

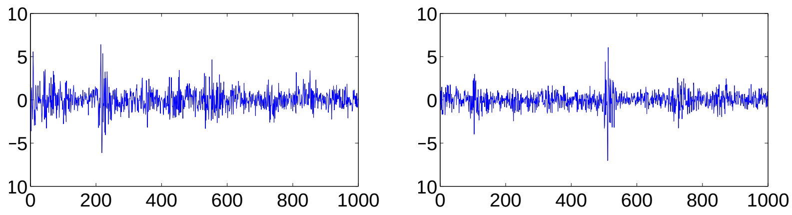
3 Wavelet-based detection procedure

3.1 A brief introduction to wavelets

Fourier analysis is a classical mathematical technique for analyzing a signal, or a time series, that transforms its view from time-based to frequency-based. It is extremely useful when the signal's frequency content is of great importance, although its main drawback is that time information is lost when transforming to the frequency domain. This drawback would be irrelevant if the series was stationary, but most interesting series contain numerous nonstationary or transitory characteristics: drift,

[1] The GARCH parameters used in the simulation were selected by fitting this model to a daily return series.

Figure 1: GARCH simulated series with an AVO outlier of (a) size $15\sigma_y$ at position 213 and (b) size $25\sigma_y$ at position 500.



(a) (b)

trends, abrupt changes and beginnings and ends of events, which are often the most important part of the series.

Wavelets are a relatively new way of analyzing time series (the formal subject dates back to 1980s) that combine old ideas with new elegant mathematical results and efficient computational algorithms. Using wavelet methods, in the context of multiresolution analysis (Mallat, 1989a,b), one can examine the series on variety of scales and different type of behavior (such as trends, cycles or extremes) may become evident at different levels of resolution. Thus wavelet analysis seems to be a useful technique to outlier detection.

A wavelet is a waveform of effectively limited duration that has an average value of zero. Compared with sine waves, which are the basis of Fourier analysis and do not have limited duration, wavelets tend to be irregular and asymmetric whereas sinusoids are smooth and predictable. Fourier analysis consists of breaking up a signal into sine waves of various frequencies. Similarly, wavelet analysis is the breaking up of the signal into shifted and scaled versions of the original (or *mother*) wavelet.

Unlike Fourier basis functions, which are only localized in frequency, wavelets are local both in frequency, via dilatations, and in time, via translations. Additionally, many classes of functions can be represented via fewer terms with wavelet transforms than with Fourier transforms. Functions with discontinuities and sharp spikes usually require fewer wavelet basis functions than Fourier basis functions. This sparse representation makes wavelets an excellent tool for data compression and statistical applications. Wavelet algorithms are fast to implement, which is especially important with large amount of data.

The algorithm we propose uses the notions of discrete wavelet transform and inverse discrete wavelet transform, hence, in the following we will only mention the concepts that are strictly necessary for the understanding of our method. We refer to Percival and Walden (2000) as a complete guide to wavelet methods for time series.

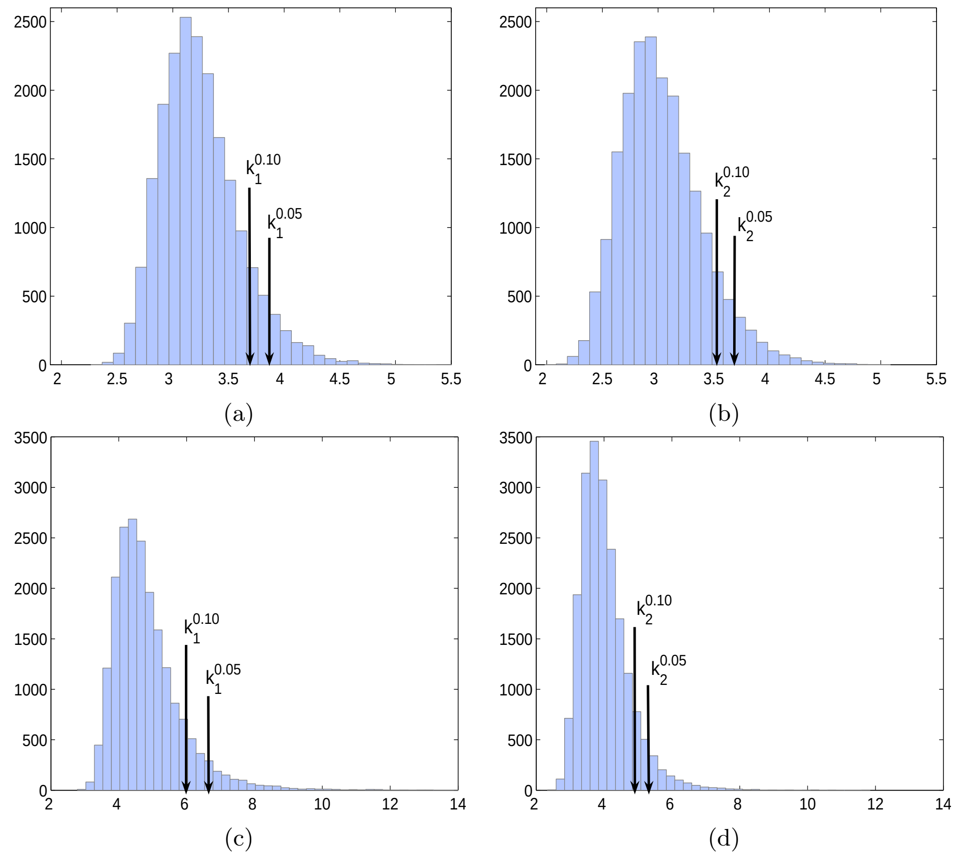
Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be the observed data (that is, the signal or the time series to be analyzed), where $X_i = f(t_i)$, $t_i = i/n$, $i = 1, \ldots, n$ and $n = 2^J$. The discrete wavelet transform (DWT) uses orthogonal transformations to decompose $\mathbf{X}$ into vectors of wavelet coefficients $\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_J$ and $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_J$, where each set of wavelet coefficients contains $n/2^j$ data points for $j = 1, \ldots, J$ (see Odgen, 1997; Cohen et al., 1993, for other methods dealing with data whose size is not a

power of two). There are many mother wavelets which can be used for computing a wavelet transform and the corresponding wavelet coefficients. In this work we use the Haar wavelet because it has good local properties and there is no need for greater regularity or smoother wavelets (see Greenblatt, 1995; Bilen and Huzurbazar, 2002). In practice, the wavelet coefficients are computed efficiently using the *pyramid algorithm*, introduced in the context of multiresolution analysis by Mallat (1989b), that is based on a pair of high and low pass filters. The low pass filter is like computing a moving average of the data, but with the difference that the weights are chosen in a very particular manner. Instead, the high pass filter consists on a moving difference of the data, and gives the detailed information. From the recursive application of the pyramid algorithm one obtains two sets of wavelet coefficients: the *approximation coefficients* $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_J$ that contain the low-frequency content, and the *detail coefficients* $\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_J$ that contain the high-frequency content. The main property of the detail coefficients is their extreme sensitivity to nonsmooth characteristics of the data such as noise, jumps and spikes. Therefore they are going to play a fundamental role in the detection of outliers.
One of the difficulties in outlier detection is the notion of *masking*, that takes place when one outlier hides others from being detected. To avoid this problem, the algorithm we propose searches recursively for only one outlier each time. This involves that, once the outlier is detected, the series is cleaned somehow, and then reconstructed for a new search. The reconstruction is performed applying the inverse discrete wavelet transform (IDWT) to the approximation coefficients and the modified detail coefficients (see Section 3.2 below for the details of the algorithm).

3.2 The procedure

The procedure we propose is based on the detail coefficients resulting from the discrete wavelet transform of the series of residuals, obtained after fitting a particular model. The outliers are identified as those observations in the original series whose detail coefficients are greater (in absolute value) than a certain threshold. We use the Haar wavelet in computing the DWT of the data since, as it was pointed out by Bilen and Huzurbazar (2002), the use of the Haar wavelet yields wavelet coefficients that are expected to be large in magnitude at times where there are jumps or outliers in the original series. In the context of financial time series, it is very common to assume that if the fitted model has captured the structure of the data, then the residuals are supposed to be independent and identically distributed (iid) random variables following either a standard normal or a Student's t distribution. Using a Monte Carlo scheme, we have obtained, for different sample sizes, the distribution of the maximum of the detail coefficients (in absolute value) resulting from the DWT of iid random variables following either a standard normal or a Student's t distribution. We depict these distributions in Figure 2 for a sample size of $n = 1000$. In practice, we have found that in order to detect isolated additive level outliers (ALOs) it is enough to work with the first level detail wavelet coefficients, that is $\mathbf{D}_1$. Using the inverse discrete wavelet transform, the procedure identifies the outliers recursively, one by one, avoiding the masking effect. However, if there are patches of ALOs or isolated additive volatility outliers (AVOs) it is necessary to use $\mathbf{D}_1$ and $\mathbf{D}_2$ to identify the influential observations. From the simulation study (see Section 3.3 below) we think that a reasonable threshold to use in the detection of isolated ALOs is the 95%-percentile of the distribution of this maximum, whereas for the detection of patches of ALOs and isolated AVOs it has been more useful the 90%-percentile. Of course this is our recommendation, but any practitioner can decide a significance level to work with.

Figure 2: Histograms for the distributions of the maximum of the absolute value of the (a) first level detail coefficients resulting from the DWT of $n = 1000$ iid $N(0,1)$ random variables, (b) second level detail coefficients resulting from the DWT of $n = 1000$ iid $N(0,1)$ random variables, (c) first level detail coefficients resulting from the DWT of $n = 1000$ iid $t(7)$ random variables, (d) second level detail coefficients resulting from the DWT of $n = 1000$ iid $t(7)$ random variables. Histograms computed from 20000 Monte Carlo samples of size n. The corresponding $j = 1, 2$ thresholds at significance level α are denoted by k_j^α.



Next we describe the steps of the procedure to detect ALOs. Let $\mathbf{X}$ be the series of residuals of size n obtained after fitting the desired model. Remember that if the model has captured the structure of the data, then the vector $\mathbf{X}$ contains n iid random variables following a standard normal distribution (or either a Student's t distribution).

Step 1 Apply the DWT to the series of residuals $\mathbf{X}$ to obtain the first level wavelet coefficients $\mathbf{A}_1 = (a_1, \ldots, a_{n/2})$ and $\mathbf{D}_1 = (d_1, \ldots, d_{n/2})$.

Step 2 Set the threshold k_1^α equal to the 95%-percentile of the distribution of the maximum of the first level detail coefficients (in absolute value) resulting from the DWT of n iid random variables following a standard normal distribution (or a Student's t distribution, if it is the case), that is $k_1^\alpha = k_1^{0.05}$.

Step 3 Find $d_{max} = \max_{1 \leq j \leq n/2}\{|d_j| > k_1^{0.05}\}$, and let s be the position of d_{max} in the vector $\mathbf{D}$.

Step 4 Set $d_{max} = 0$ and construct $\widetilde{\mathbf{D}}_1$ as the vector equal to $\mathbf{D}_1$ but with a 0 in the s position, that is $\widetilde{\mathbf{D}}_1 = (d_1, \ldots, d_{s-1}, 0, d_{s+1}, \ldots, d_{n/2})$.

Step 5 Recompose the series of residuals applying the inverse discrete wavelet transform (IDWT) to $\mathbf{A}_1$ and $\widetilde{\mathbf{D}}_1$.

Step 6 Repeat steps 1 to 5 until all the elements in the vector of the detail coefficients are lower (in absolute value) than the threshold $k_1^{0.05}$. Let $S = \{s_1, \ldots, s_\ell\}$ be the ordered set of indices containing the positions of the d_{max}'s.

Step 7 Use S to locate the exact positions of the outliers in the series of residuals $\mathbf{X}$. Let s be a generic element in S. Compute the sample mean of $\mathbf{X}$ without observations at locations $2s$ and $2s-1$:

$$\overline{x}_{n-2} = \frac{1}{n-2} \sum_{i \neq 2s, 2s-1} X_i$$

and set the position of the outlier equal to $2s$ if $|X_{2s} - \overline{x}_{n-2}| > |X_{2s-1} - \overline{x}_{n-2}|$, or equal to $2s-1$, otherwise.

The algorithms that respectively search for patches of ALOs and for AVOs differ from the previous one in the sense that two level wavelet coefficients are computed and, consequently, there are two thresholds $k_1^{0.10}$ and $k_2^{0.10}$, one for each set of detail wavelet coefficients $\mathbf{D}_1$ and $\mathbf{D}_2$. But the main idea remains unchanged. These algorithms have been implemented in Matlab and are available from the authors upon request.

3.3 Performance of the procedure: A simulation study

In this Section we present the results of a simulation study to asses the performance of the proposed wavelet-based detection procedure. We have compared our method with the one proposed by Bilen and Huzurbazar (2002), which is also based on wavelets. The measures used in the performance study are the proportion of times that the location of the outliers is correctly detected jointly with the average number of false detections and their standard errors. The study involves single, multiple and patches of additive level outliers observed in different financial models, such as GARCH, GJR and ARSV with errors following either a gaussian or a Student's t distribution. We have considered magnitudes of isolated ALOs of $\omega_{AO} = 5\sigma_y, 10\sigma_y, 15\sigma_y$ and sample sizes of $n = 500, 1000, 5000$. The frequency of the simulations is daily and the parameters used are: $\{\alpha_0 = 0.0126, \alpha_1 = 0.0757, \beta = 0.9122\}$ for the GARCH model, $\{\alpha_0 = 0.0000, \alpha_1 = 0.0139, \beta = 0.9139, \gamma_1 = 0.1106\}$ for the GJR model and $\{\phi = 0.98, \sigma_\eta^2 = 0.05, \sigma = 1\}$ for the ARSV model, which have been chosen by fitting the models to real return series. Concerning additive volatility outliers, we have simulated single AVOs of magnitude $\omega_{AO} = 15\sigma_y, 25\sigma_y, 50\sigma_y$ in GARCH and GJR models with errors following either a gaussian or a Student's t distribution for a sample size of $n = 1000$. The outliers are placed randomly along the time series. The simulation study involves 1000 replications for each scenario. The threshold values used in the algorithm are presented in Table 1. They have been computed from 20000 Monte Carlo samples as the $(1-\alpha)100\%$-percentiles of the distribution of the maximum of the absolute values of the j-th level detail coefficients $(j = 1, 2)$ resulting from the DWT of n iid random variables following a standard normal or a Student's t distribution.

Table 1: Threshold values: $(1-\alpha)100\%$-percentiles of the distribution of the maximum of absolute values of the j-th level detail coefficients $(j = 1, 2)$.

		$N(0,1)$		$t(7)$	
	n	k_1^α	k_2^α	k_1^α	k_2^α
$\alpha = 0.05$	500	3.7216	3.5280	6.0053	4.9542
	1000	3.8965	3.7114	6.6477	5.3078
	5000	4.2620	4.0992	8.2632	6.4090
$\alpha = 0.10$	500	3.5273	3.3339	5.4636	4.5236
	1000	3.7104	3.5277	6.0253	4.9052
	5000	4.0935	3.9285	7.5061	5.9045

Tables 2 and 3 contain the results for the single and multiple ALOs. These outliers have been detected using the corresponding $k_1^{0.05}$ threshold value reported in Table 1. When the magnitude of the outliers is $\omega_{AO} = 10\sigma_y, 15\sigma_y$, the procedure detects more than 90% of single and multiple outliers, for models with gaussian errors. When the errors follow a Student's t distribution, the detection rate goes from 52% to 95%, being around 80% in mean. As expected, the sensitivity of the procedure increases as the magnitude of the outlier increases, specially in the case the errors follow a Student's t distribution because small size outliers can not be distinguished from the thick tail of Student's $t(7)$ distribution. Additionally, the average number of false detections is no greater than 0.65 (note that it is no greater than 0.1 in practically all cases). This fact makes the procedure very reliable, in the sense that one can be sure that any observation identified as an outlier it is an outlier. It is true that the detection rate of the method of Bilen and Huzurbazar (2002) is slightly greater than ours, specially for outliers of small size, but it is also true that their average number of false detections is, in almost all cases, extremely high compared to ours and their methodology is only valid assuming the data generating process is composed of normal random variables. Moreover, we observe from the outlier detection results that the ARSV and GJR models are more robust to outliers of small size in the sense that they can not be distinguished from the observations generated by the two specifications.

We have also compared the effectiveness of the presented methods with what we have called the *naive method*, which is a commonly used practise for its simplicity and applicability. The *naive method* classifies as an outlier a residual observation that exceeds seven residual standard deviations. Table 4 reports these comparisons

Table 2: Percentage of correct detection of additive level outliers in 1000 replications of size n for various volatility models with errors following a normal or a Student's t distribution.

		$N(0,1)$						$t(7)$		
		GARCH		ARSV		GJR		GARCH	ARSV	GJR
	n	G&V	B&H	G&V	B&H	G&V	B&H			
1 outlier	500	66.3	93.4	64.1	87.1	25.8	95.5	39.9	17.1	30.3
of size	1000	66.0	90.2	63.6	85.3	20.7	92.6	41.9	17.5	37.1
$\omega_{AO}=5$	5000	59.6	86.1	60.5	80.1	10.7	93.1	88.5	9.1	56.9
1 outlier	500	98.4	100	96.9	99.7	92.1	98.5	68.1	68.5	78.4
of size	1000	98.9	99.9	96.0	99.0	92.9	99.0	70.3	66.1	77.9
$\omega_{AO}=10$	5000	98.4	99.7	94.0	97.6	97.1	99.8	93.5	52.5	78.2
1 outlier	500	99.0	100	99.5	99.9	91.0	97.9	79.3	93.1	92.2
of size	1000	99.7	100	99.6	100	94.8	98.3	79.8	88.5	91.4
$\omega_{AO}=15$	5000	99.8	99.9	98.6	99.8	99.6	100	95.8	80.7	87.5
3 outliers	500	63.3	91.8	71.3	95.2	63.7	92.5	35.9	40.8	51.8
of sizes	1000	71.4	92.0	76.9	94.9	64.9	92.1	48.8	47.6	61.3
$\omega_{AO}=5,10,15$	5000	77.6	90.3	81.3	92.5	65.1	91.9	80.9	45.7	73.2

Table 3: Average number of false detections (standard deviation) of additive level outliers in 1000 replications of size n for various volatility models with errors following a normal or a Student's t distribution.

		$N(0,1)$						$t(7)$		
		GARCH		ARSV		GJR		GARCH	ARSV	GJR
	n	G&V	B&H	G&V	B&H	G&V	B&H			
1 outlier	500	0.02	1.96	0.14	2.50	0.02	3.64	0.001	0.01	0.01
of size		(0.15)	(7.99)	(0.37)	(1.97)	(0.13)	(2.77)	(0.03)	(0.11)	(0.11)
$\omega_{AO}=5$	1000	0.05	1.91	0.24	3.43	0.01	5.02	0.01	0.02	0.02
		(0.22)	(1.58)	(0.50)	(2.20)	(0.11)	(2.91)	(0.10)	(0.13)	(0.14)
	5000	0.05	2.63	0.65	7.17	0.01	10.52	0.03	0.03	0.02
		(0.21)	(1.73)	(0.78)	(2.92)	(0.11)	(3.74)	(0.18)	(0.18)	(0.12)
1 outlier	500	0.03	3.85	0.09	2.52	0.03	3.92	0.01	0.01	0.01
of size		(0.20)	(19.25)	(0.30)	(1.97)	(0.17)	(2.90)	(0.08)	(0.08)	(0.12)
$\omega_{AO}=10$	1000	0.03	2.21	0.15	3.31	0.02	5.15	0.01	0.01	0.02
		(0.16)	(1.91)	(0.40)	(2.15)	(0.13)	(3.00)	(0.10)	(0.11)	(0.13)
	5000	0.04	2.71	0.57	7.16	0.01	10.54	0.03	0.03	0.02
		(0.19)	(1.82)	(0.73)	(2.90)	(0.11)	(3.79)	(0.17)	(0.17)	(0.12)
1 outlier	500	0.04	5.07	0.04	2.51	0.06	4.32	0.01	0.005	0.01
of size		(0.20)	(22.16)	(0.21)	(2.03)	(0.26)	(3.34)	(0.08)	(0.07)	(0.11)
$\omega_{AO}=15$	1000	0.04	4.35	0.10	3.41	0.03	5.45	0.01	0.01	0.02
		(0.21)	(27.28)	(0.31)	(2.20)	(0.18)	(3.17)	(0.10)	(0.08)	(0.13)
	5000	0.03	2.84	0.49	7.32	0.01	10.50	0.03	0.02	0.01
		(0.17)	(1.96)	(0.71)	(2.96)	(0.12)	(3.83)	(0.17)	(0.13)	(0.12)
3 outliers	500	0.03	5.00	0.02	2.72	0.09	5.10	0.001	0.002	0.01
of sizes		(0.19)	(8.84)	(0.15)	(2.10)	(0.33)	(4.45)	(0.03)	(0.04)	(0.11)
$\omega_{AO}=5,10$	1000	0.04	7.34	0.04	3.25	0.07	5.80	0.004	0.003	0.01
and 15		(0.24)	(41.28)	(0.22)	(2.11)	(0.29)	(3.42)	(0.06)	(0.05)	(0.11)
	5000	0.03	3.15	0.35	7.14	0.02	10.70	0.04	0.01	0.01
		(0.17)	(2.09)	(0.62)	(3.00)	(0.12)	(4.03)	(0.19)	(0.12)	(0.11)

Table 4: Percentage of correct detection and average number of false detections (standard deviation), with respect to the *naive method*, of additive level outliers of size $\omega_{AO} = 5$ in 1000 replications of size n for various volatility models with errors following a normal or a Student's t distribution.

	$N(0,1)$						$t(7)$		
	GARCH		ARSV		GJR		GARCH	ARSV	GJR
n	G&V	B&H	G&V	B&H	G&V	B&H			
500	100	100	100	100	90.3	100	50.8	17.8	61.7
1000	100	100	99.4	100	77.5	100	59.7	18.7	58.9
5000	99.7	100	98.0	100	58.8	100	77.0	12.2	56.2

Percentage of correct detection of outliers

	$N(0,1)$						$t(7)$		
	GARCH		ARSV		GJR		GARCH	ARSV	GJR
n	G&V	B&H	G&V	B&H	G&V	B&H			
500	0.59	2.80	0.49	3.07	0.22	4.53	0.001	0.01	0.01
	(0.52)	(8.01)	(0.60)	(2.05)	(0.42)	(2.79)	(0.03)	(0.07)	(0.10)
1000	0.51	2.61	0.50	3.91	0.15	5.86	0.001	0.004	0.002
	(0.54)	(1.63)	(0.68)	(2.27)	(0.37)	(2.95)	(0.03)	(0.06)	(0.04)
5000	0.31	3.16	0.80	7.51	0.06	11.35	0.02	0.001	0
	(0.48)	(1.77)	(0.87)	(2.97)	(0.24)	(3.78)	(0.15)	(0.03)	-

Average number of false detections (standard deviation) of outliers

from where we observe that, in some cases, the outlier detection method proposed by Bilen and Huzurbazar (2002) detects more than ours but, in all situations, their average number of false detections is also much higher. Note that for the GJR model with gaussian errors, our method detects 58.5% of the outliers detected by the *naive method* versus the 100% detection rate of Bilen and Huzurbazar (2002)'s, but their average of false detections is nearly 190 times greater than ours.

We have also studied the effectiveness of our procedure in detecting patches of additive level outliers. We have considered patches of three ALOs in the same volatility models as before (GARCH, GJR and ARSV with errors following either a gaussian or a Student's t distribution). We have taken magnitudes of $\omega_{A0} = 10\sigma_y, 15\sigma_y$ and sample sizes of $n = 500, 1000, 5000$. The beginning of the patch was placed randomly in the time series. The simulation study involves 1000 replications for each scenario. Tables 5 and 6 contain the results for the patches of ALOs. These outliers have been detected using the corresponding $k_1^{0.10}$ and $k_2^{0.10}$ threshold values reported in Table 1. In general, the detection rate is greater for models with gaussian errors, going from 41% to nearly 98%, whereas the average number of false detections is always no greater than 0.03. The sensitivity of the method increases as ω_{AO} increases.

Finally, we have studied the effectiveness of our method in detecting single additive volatility outliers. We have considered AVOs in the GARCH and GJR models (with errors following either a gaussian or a Student's t distribution). We have taken magnitudes of $\omega_{AO} = 15\sigma_y, 25\sigma_y, 50\sigma_y$ and a sample size of $n = 1000$. As before, the AVO was placed randomly in the time series. The simulation study involves 1000 replications for each scenario. Table 7 contains the results for the AVO's detection. These outliers have been detected using the corresponding $k_1^{0.10}$ and $k_2^{0.10}$ threshold

Table 5: Percentage of correct detection of patches of 3 additive level outliers in 1000 replications of size n for various volatility models with errors following a normal or a Student's t distribution.

		$N(0,1)$			$t(7)$		
	n	GARCH	ARSV	GJR	GARCH	ARSV	GJR
1 patch of 3	500	82.1	59.8	76.6	52.5	21.6	58.3
outliers of	1000	78.5	66.1	69.5	61.9	25.2	62.4
size $\omega_{AO} = 10$	5000	74.5	70.9	41.1	93.0	26.6	70.3
1 patch of 3	500	97.8	86.0	87.1	73.3	48.0	80.2
outliers of	1000	97.7	92.1	87.8	72.9	57.0	80.6
size $\omega_{AO} = 15$	5000	96.1	90.4	92.7	95.6	54.3	81.5

Table 6: Average number of false detections (standard deviation) of patches of 3 additive level outliers in 1000 replications of size n for various volatility models with errors following a normal or a Student's t distribution.

		$N(0,1)$			$t(7)$		
		GARCH	ARSV	GJR	GARCH	ARSV	GJR
1 patch of 3	500	0.001	0	0	0.005	0.005	0.004
outliers of		(0.03)	-	-	(0.07)	(0.07)	(0.06)
size $\omega_{AO} = 10$	1000	0.001	0.01	0	0.005	0.01	0.01
		(0.03)	(0.08)	-	(0.07)	(0.08)	(0.08)
	5000	0	0.02	0	0.03	0.02	0.01
		-	(0.14)	-	(0.17)	(0.13)	(0.09)
1 patch of 3	500	0	0.001	0.002	0.005	0.002	0.01
outliers of		-	(0.03)	(0.04)	(0.07)	(0.05)	(0.08)
size $\omega_{AO} = 15$	1000	0.001	0.001	0	0.005	0.005	0.01
		(0.03)	(0.03)	-	(0.07)	(0.07)	(0.08)
	5000	0	0.01	0	0.03	0.01	0.01
		-	(0.10)	-	(0.17)	(0.09)	(0.11)

values reported in Table 1. In general, the detection rate is greater for models with gaussian errors, whereas the average number of false detections is always no greater than 0.004. Again, the sensitivity of the method increases with ω_{AO}. In the case of AVO outliers, we also observe that the method detects more outliers in the GJR model than in the GARCH model. This finding can be justified by the fact that an AVO of the same size may have a more amplified impact in the GJR model than in the GARCH model, depending on the parameters chosen to simulate the respective models (compare the volatility equations 5 and 6 of both specifications).

4 Empirical Applications

In this Section we analyze three financial return series to illustrate the performance of our method on real data. The series we consider are the Dow Jones, the FTSE-100 and the S&P 500 indexes. The data was collected from Yahoo Finance website (`http://finance.yahoo.com/`) and spans the period of April 2, 1984-July 29, 2008. Figure 3 depicts the three return series, $y_t = (\log p_t - \log p_{t-1}) \cdot 100$, where p_t is the value at time t of the corresponding index and Table 8 reports some descriptive

Table 7: Percentage of correct detection and average number of false detections of additive volatility outliers in 1000 replications of size $n = 1000$ for two volatility models with errors following a normal or a Student's t distribution.

	$N(0,1)$		$t(7)$	
w	GARCH	GJR	GARCH	GJR
15	36.2	94.6	30.4	77.1
25	54.7	96.6	48.8	81.1
50	71.4	96.5	65.8	83.8

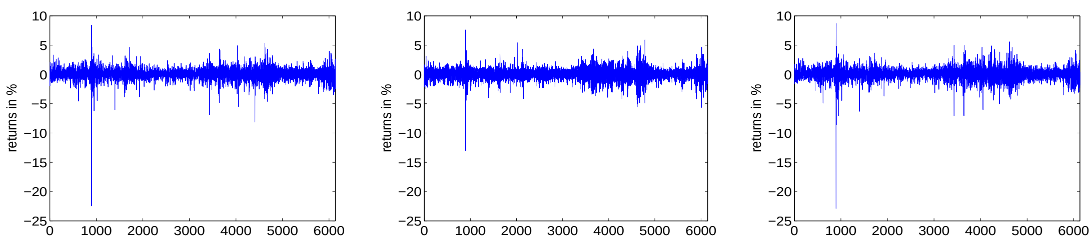
Percentage of correct detection of AVOs

	$N(0,1)$		$t(7)$	
w	GARCH	GJR	GARCH	GJR
15	0.004	0.003	0	0
	(0.063)	(0.055)	-	-
25	0.001	0	0	0
	(0.032)	-	-	-
50	0	0.001	0	0
	-	(0.113)	-	-

Average number of false detections (standard deviation) of AVOs

statistics. From Table 8, we observe that the three return series are negatively skewed and have significant kurtosis, ranging from 10.771 for the FTSE-100 to 54.929 for the Dow Jones, which make us suspect of the existence of some outliers. It is known that the existence of extreme observations in time series leads to fat tail distributions and some outlier detection methods, specially in the multivariate context, are based on this information (see for example Peña and Prieto, 2001; Galeano et al., 2006). Table 8 also contains the results of the Kiefer and Salmon (1983) test, that is a formal test of normality in the context of conditional heteroscedastic series.[2]

Figure 3: Returns in percentage for (a) Dow Jones index, (b) FTSE-100 index and (c) S&P 500 index.



(a) (b) (c)

[2]The Kiefer and Salmon (1983) test is given by $KS_N = (KS_S)^2 + (KS_K)^2$, where $KS_S = \sqrt{\frac{T}{6}}\left[\frac{1}{T}\sum_{t=1}^{T} y_t^{*3} - \frac{3}{T}\sum_{t=1}^{T} y_t^{*}\right]$, $KS_K = \sqrt{\frac{T}{24}}\left[\frac{1}{T}\sum_{t=1}^{T} y_t^{*4} - \frac{6}{T}\sum_{t=1}^{T} y_t^{*2} + 3\right]$ and y_t^* are the standardized returns. If the distribution of y_t^* is conditional $N(0,1)$ then KS_S and KS_K are asymptotically $N(0,1)$ and KS_N is asymptotically $\chi^2(2)$.

Table 8: Descriptive statistics for the daily stock index returns.

Stock index returns	Dow Jones	FTSE-100	S&P 500
Mean	0.038	0.028	0.040
Variance	1.043	1.125	1.194
Skewness	-2.413^*	-0.539^*	-2.035^*
Kurtosis	54.929^*	10.771^*	46.107^*
KS_S	-67.963	-16.338	-59.646
KS_K	744.493	117.877	653.651

In order to check for the presence of outliers, we have applied the proposed wavelet-based procedure to the residual series. We decided to apply our procedure to detect only isolated and patches of ALO outliers since the effects of a volatility outlier on the level series are similar to a patch of outliers with decreasing magnitudes over time. Table 9 contains the results of the isolated ALO detection, using a threshold value of $k_1^{0.05} = 4.3042$ computed from 20000 Monte Carlo samples of size $n = 6100$. The observation positions presented in Table 9 are already the ones of the real data. We observe that our procedure only detects outliers in the gaussian series of residuals (see also Figures 5, 6 and 7). In fact, the Student's t volatility models seem to capture properly the data features. We also observe for the Dow Jones return series that observations 896, 1399 and 3431 are considered ALO outliers in all models. The first observation corresponds to October 19, 1987. The day that subsequently became known as "Black Monday". Both, the Dow Jones and the S&P 500 lost more than twenty percent of their total value in that day. The second outlier detected corresponds to October 13, 1989. This day corresponded also to a crash that was apparently caused by a reaction to a news story of a \$6.75 billion leveraged buyout deal for UAL Corporation. The parent company of United Airlines, which fell through. Finally, the observation 3431 corresponds to October 27, 1997, a mini crash caused by an economic crisis in Asia. These same observations are also detected as ALO outliers in the S&P 500 residuals for all models. Some other observations are also considered ALO outliers for the Dow Jones and GARCH-type models. For instance, observation 4407 (September 17, 2001) that corresponds to the first day open of the New York Stock Exchange after the terrorist attack to the USA on the 11th of September, 2001. Regarding the S&P 500 there is also another observation (observation 5777) that is detected as an ALO outlier for the considered models. It corresponds to February 27, 2007, the day of the big decline in Chinese stocks and the news of the weakness in some key readings on the U.S. economy. Regarding the FTSE-100, we also observe that observations 4785 and 4786 correspond to March 19, 2003 and March 20, 2003, respectively. These days corresponded to the reaction of share prices in London to the expected onset of hostilities in the Gulf. Observation 3980 (December 31, 1999) is also considered an ALO outliers for the GARCH-type models. It corresponded to a market correction since on December 30, 1999 the FTSE-100 reached its highest value to the date. Our procedure is quite effective in capturing the most important crashes in three important international stock markets, the New York Stock Exchange, NASDAQ and the London Stock Exchange.

Figure 4 shows the graphical output of the Matlab program, which corresponds to

Table 9: Observations identified as possible additive level outliers for $\alpha = 0.05$ in the three series of stock market indexes.

Dow Jones

GARCH		GJR		ARSV	
$N(0,1)$	$t(7)$	$N(0,1)$	$t(7)$	$N(0,1)$	$t(7)$
896	-	896	-	896	-
1399		1399		1399	
1928		1928		2530	
3431		3431		3431	
4407		4053			
		4407			

FTSE-100

GARCH		GJR		ARSV	
$N(0,1)$	$t(7)$	$N(0,1)$	$t(7)$	$N(0,1)$	$t(7)$
3980	-	3980	-	897	-
4785				4786	

S&P 500

GARCH		GJR		ARSV	
$N(0,1)$	$t(7)$	$N(0,1)$	$t(7)$	$N(0,1)$	$t(7)$
896	-	1399	-	896	-
1399		3431		901	
3431		5777		1399	
3643				2530	
5777				3431	
				3643	
				5777	

Threshold values: $k_1^{0.05} = 4.3042$ for $N(0,1)$, $k_1^{0.05} = 8.5091$ for $t(7)$.

the analysis of the S&P 500 residuals obtained from a GARCH model with gaussian errors.

Figure 4: Graphical output of the wavelet-based procedure for the returns of S&P 500 estimated as a GARCH model with gaussian errors.

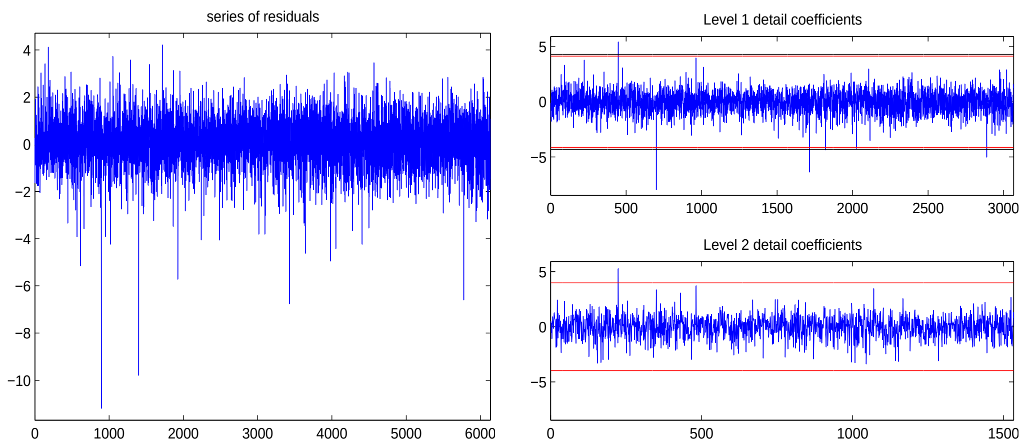


Figure 5: Dow Jones residuals.

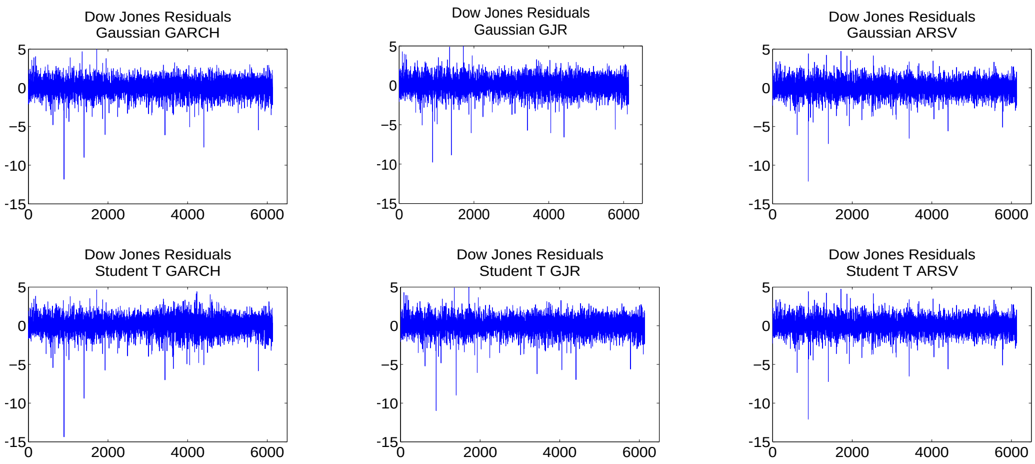
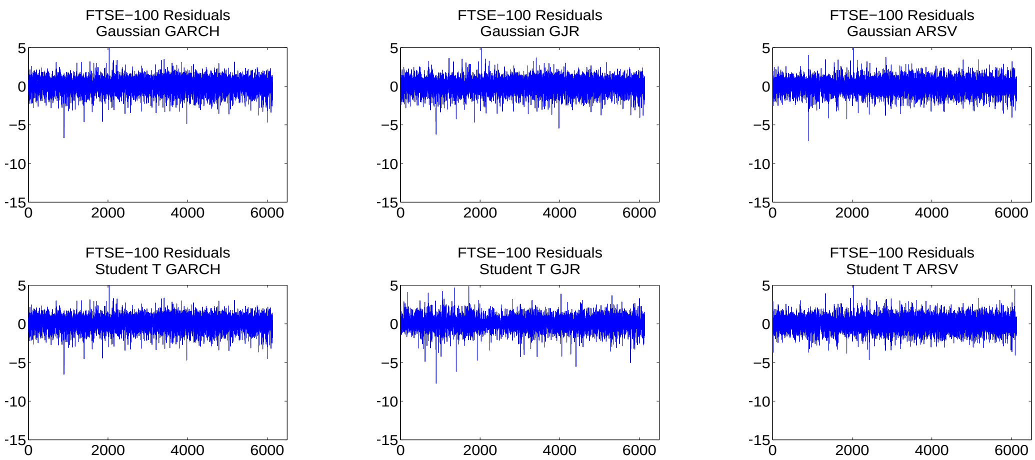
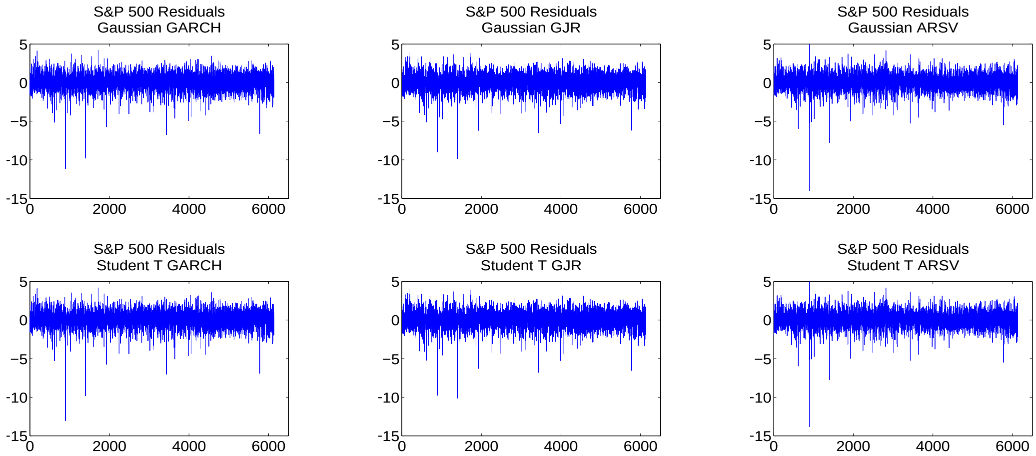


Figure 6: FTSE-100 residuals.



Regarding ALO patches, we have used threshold values of $k_1^{0.10} = 4.1476$ and $k_2^{0.10} = 3.9738$ computed from 20000 Monte Carlo samples of size $n = 6100$. We have detected one patch (around observation 896) for both Dow Jones and S&P 500 residual series obtained with a gaussian GARCH. This observation was also detected as an isolated ALO outlier. In fact, this patch corresponds to the days around the stock market crash of October 19, 1987. Additionally, we have also found a patch around 1928 (November 15, 1991) in the S&P 500 residual series obtained with a GJR model with gaussian errors. For other considered scenarios (other series and/or other type of errors) the algorithm detects no more patches.

Figure 7: S&P 500 residuals.



4.1 Correction of the detected outliers

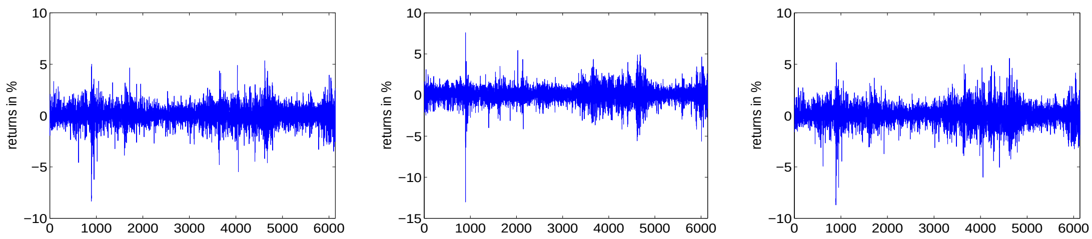
Once the positions of the outliers in the series have been determined, we propose to remove those abnormal observations from the series using the same methodology as in the algorithm described in Section 3.2. The idea consists in decomposing the series using the discrete wavelet transform in order to obtain the first and second level wavelet coefficients. Let $\{\mathbf{A}_1, \mathbf{D}_1\}$ and $\{\mathbf{A}_2, \mathbf{D}_2\}$ denote the pairs of vectors containing, respectively, those first and second level approximation and detail wavelet coefficients.

(a) Correcting for isolated ALOs: Assign zero to those elements in $\mathbf{D}_1$ that have been used in the detection process to identify isolated ALOs and denote by $\widetilde{\mathbf{D}}_1$ the corrected first level detail coefficients. To reconstruct the series, apply the inverse wavelet transform to $\mathbf{A}_1$ and $\widetilde{\mathbf{D}}_1$.

(b) Correcting for patches of ALOs: Assign zero to those elements in $\mathbf{D}_2$ and $\mathbf{D}_1$ that have been used in the detection process to identify a patch of ALOs. Denote by $\widetilde{\mathbf{D}}_2$ and $\widetilde{\mathbf{D}}_1$ the corrected second and first level detail coefficients, respectively. Apply the inverse wavelet transform to $\mathbf{A}_2$ and $\widetilde{\mathbf{D}}_2$ to reconstruct the first level approximation coefficients and denote them by $\widetilde{\mathbf{A}}_1$. To reconstruct the series apply the inverse wavelet transform to $\widetilde{\mathbf{A}}_1$ and $\widetilde{\mathbf{D}}_1$.

We use the algorithm previously described to correct the outliers found in the three financial return series when fitting a GARCH model. Figure 8 contains these "corrected" return series (in percentage).
Table 10 reports summary statistics of the return series corrected for the outliers detected using the traditional GARCH model. As it was expected, we observe that the kurtosis of the cleaned return series decreased in all series and in particular, for the Dow Jones and S&P 500 data.

Figure 8: Return series in percentage, after correcting the possible outliers, for (a) Dow Jones index, (b) FTSE-100 index and (c) S&P 500 index.



(a) (b) (c)

Table 10: Descriptive statistics for the daily stock index returns corrected for the outliers detected in the gaussian GARCH residuals.

Stock index returns	Dow Jones	FTSE-100	S&P 500
Mean	0.036	0.026	0.035
Variance	0.918	1.064	1.015
Skewness	-0.498^*	-0.523^*	-0.452
Kurtosis	8.974^*	10.309^*	8.878^*
KS_S	-15.920	-16.739	-14.437
KS_K	95.4804	116.9050	93.9476

5 Conclusion

The existing outlier procedures in financial time series are based on the proposal by Chen and Miu (1993b) that consists in an iterative outlier detection and adjustment method to jointly estimate the model parameters and the outlier effects. However, along the iterative process they have to estimate the model several times and the estimates of the parameters can be affected by the presence of remaining outliers.
On the contrary, our outlier detection proposal is based on applying wavelets to the residuals of some volatility models. It does not need subsequential estimations of the model parameters and therefore, it is not susceptible to the previous criticism. The method uses the discrete wavelet transform and detects changes in the wavelet coefficients by using thresholds based on the distribution of the maximum of the detailed coefficients (in absolute value) obtained by Monte Carlo. In this way, our method can be applied to the estimated residuals of different volatility models with errors following any known distribution.
The advantages of our proposal are several: first, it applies when the location and the number of outliers are unknown, second, the data can be generated by any known distribution; third, it is well suited for one outlier or multiple outlier detections; fourth, it is the one, as far as we know, that detects patches of outliers in different volatility models; and finally, the method is easy and quick to apply which converts it in an attractive tool to be used by academic communities and/or by practitioners.
The effectiveness of our method is tested both with simulated and real data and

it is compared with other outlier detection methods. The simulations report evidence that our proposal is not only as good as that of Bilen and Huzurbazar (2002), whenever both methods can be applied, but also much reliable since it detects a significant smaller number of false outliers. Moreover, since Bilen and Huzurbazar (2002) showed that their outlier detection procedure performed better that the ones based on likelihood ratio tests like the method by Chen and Miu (1993b), we may conclude that our detection method is better than the existing proposals in financial time series, with the advantage that we can test for patches of additive level outliers and data generated from different known distributions.

References

Baillie, R. and T. Bollerslev (1989). The message in daily exchange rates: a conditional variance tale. *Journal of Business and Economic Statistics 7*, 297–309.

Bilen, C. and S. Huzurbazar (2002). Wavelet-based detection of outliers in time series. *Journal of Computational and Graphical Statistics 11*, 311–327.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economic and Statistics 69*, 542–547.

Carnero, M., D. Peña, and E. Ruiz (2007). Effects of outliers on the identification and estimation of GARCH models. *Journal of Time Series Analysis 28*, 471–497.

Chen, C. and L. Miu (1993a). Forecasting time series with outliers. *Journal of Forecasting 12*, 13–35.

Chen, C. and L. Miu (1993b). Joint estimation of model parameters and outlier effects. *Journal of American Statistical Association 88*, 284–297.

Cohen, A., I. Daubechies, B. Jawerth, and P. Vial (1993). Multiresolution analysis, wavelets and fast algorithms on an interval. *Comptes Rendus de l'Académie des Sciences, Série I 316*, 417–421.

Donoho, D. and I. Johnstone (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika 81*, 425–455.

Doornik, J. and M. Ooms (2005). Outlier detection in GARCH models. *Tinbergen Institute Discussion Paper*.

Duan, J.-C., G. Gauthier, J.-G. Simonato, and C. Sasseville (2006). Approximating the GJR-GARCH and EGARCH option pricing models analytically. *Journal of Computational Finance 9*, 41–69.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica 50*, 987–1008.

Fox, A. (1972). Outliers in time series. *Journal of Royal Statistical Society B 34*, 350–363.

Franses, P. and H. Ghijsels (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting 15*, 1–9.

Galeano, P., D. Peña, and R. Tsay (2006). Outlier detection in multivariate time series by projection pursuit. *Journal of the American Statistical Association 101*, 654–669.

Glosten, L., R. Jagannathan, and D. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance 48*, 1779–1801.

Greenblatt, S. A. (1995). Wavelets in econometrics: An application to outlier testing. In M. Gilli (Ed.), *Computational Economic Systems: Models, Methods, and Econometrics*, Advances in Computational Economics, 5, pp. 139–160. EKluwer Academic Publishers, Dordrecht.

Hotta, L. and R. Tsay (1998). Outliers in GARCH processes. *Manuscript. Graduate School of Business, University of Chicago.*

Kiefer, N. and M. Salmon (1983). Testing normality in econometric models. *Economics Letters 11*, 123–127.

Ledolter, J. (1989). The effect of additive outliers on the forecasts from ARIMA models. *International Journal of Forecasting 5*, 231–240.

Mallat, S. (1989a). Multiresolution approximations and wavelet orhtonormal bases of $L^2(\mathbb{R})$. *Transsactions of the American Mathematical Society 315*, 69–87.

Mallat, S. (1989b). A theory for multiresolution signal decomposition: the wavelet representation. IEEE *Pattern Analalysis and Machine Intelligence 11*, 674–693.

Odgen, R. (1997). *Essential Wavelets for Statistical Applications and Data Analysis.* Boston: Birkhäuser.

Peña, D. (2001). Outliers, influential observations and missing data. In D. Peña, G. Tiao, and R. Tsay (Eds.), *A Course in Time Series*, New York, pp. 136–170. Wiley.

Peña, D. and F. Prieto (2001). Multivariate outlier detection and robust covariance matrix estimation. *Technometrics 43*, 286–310.

Percival, D. and A. Walden (2000). *Wavelet Methods for Time Series Analysis.* New York: Cambridge University Press.

Sakata, S. and H. White (1998). High breakdown point conditional dispersion estimation with aplication to S&P500 daily returns volatility. *Econometrica 66*, 529–567.

Taylor, S. (1986). *Modelling Financial Time Series.* Wiley, New York.

Teräsvirta, T. (1996). Two stylized facts and the GARCH(1,1) model. *Working Paper 96, Stockholm School of Economics*.

Van Dijk, D., P. Franses, and A. Lucas (1999). Testing for ARCH in the presence of additive outliers. *Journal of Applied Econometrics 14*, 539–562.

Wang, Y. (1995). Jump and sharp cusp detection by wavelets. *Biometrika 82*, 385–397.